DOCUMENT RESUME

ED 084 492 CS 000 816

AUTHOR Hiller, Jack H.
TITLE Effectiveness of Various Prose Study Techniques as a Function of Text Readability, and Individual Differences.
PUB DATE 73
NOTE 41p.; Unpublished report of research conducted at the Southwest Regional Laboratory for Educational Research and Development

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS College Students; Learning; Memory; *Questioning Techniques; *Readability; Reading; *Reading Research; Reading Skills; *Recall (Psychological); Retention; *Retention Studies; Study Skills

ABSTRACT

Immediate and two-week retention were studied as a function of three levels of text readability and two levels of inserted post question (IPQ) difficulty. The IPQ treatment was modified to permit review of text after question answering. A traditional control group was required to read without marking lesson pages; and a second control was permitted complete freedom to study. Both IPQ treatments produced significantly inferior acquisition of content incidental to the IPQs for the two lessons having below average readability. For the average readability lesson, only the difficult IPQs produced significantly lower acquisition. Treatment differences diminished to non-significant levels on two-week retention. Learning decrement was correlated with test anxiety and self-confidence in the difficult lessons but not in the average readability lesson. (Author)

FILMED FROM BEST AVAILABLE COPY

ED 084492

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

EFFECTIVENESS OF VARIOUS PROSE STUDY TECHNIQUES AS A FUNCTION OF TEXT READABILITY, AND INDIVIDUAL DIFFERENCES

Jack H. Hiller[1]

Southwest Regional Laboratory for

Educational Research and Development

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY

Jack H. Hiller

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

reduced version to be published in J. of Educ Psych

CS 000 816

EFFECTIVENESS OF VARIOUS PROSE STUDY TECHNIQUES AS A FUNCTION OF TEXT READABILITY, AND INDIVIDUAL DIFFERENCES

ABSTRACT

Immediate and two-week retention were studied as a function of three levels of text readability and two levels of inserted post question (IPQ) difficulty. The IPQ treatment was modified to permit review of text after question answering. A traditional control group was required to read without marking lesson pages; and a second control was permitted complete freedom to study. Both IPQ treatments produced significantly inferior acquisition of content incidental to the IPQs for the two lessons having below average readability. For the average readability lesson, only the difficult IPQs produced significantly lower acquisition. Treatment differences diminished to non-significant levels on two-week retention. Learning decrement was correlated with test anxiety and self-confidence in the difficult lessons but not in the average readability lesson.

EFFECTIVENESS OF VARIOUS PROSE STUDY TECHNIQUES AS A FUNCTION OF TEXT READABILITY, AND INDIVIDUAL DIFFERENCES

Numerous studies have investigated the effects of adjunct questions on learning from prose text. A set of widely held conclusions is that inserted post-questions raise the learning of: a) specific details asked about by the questions (Rothkopf, 1966); b) details categorically related to the questions (Rothkopf & Bisbicos, 1967); and c) information not directly related to the questions (Rothkopf, 1966; Frase, 1968). Where a lesson contains more information than a student can learn during a given study period, specific questions surely enhance question related learning by directing and focusing attention (Quellmalz, 1971) and by stimulating rehearsal (McGaw & Grotelueschen, 1972). The intriguing experimental finding is that incidental learning is not reduced as a consequence of inserted post-question treatments; therefore, incidental learning alone was investigated in this study.

Rothkopf (1970) and Frase (1970) have cautiously concluded that inserted post-questions may enhance learning under circumstances that are not yet well defined nor understood. Possible bases for the effects of adjunct questions have been suggested by Frase (1970): "Questions are motivational stimuli. They have arousal and associative outcomes," (page 346). Arousal may be related to drive level, and theorists have long held that the relationship between learning-performance and drive is curvilinear, and a function of task complexity (Yerkes & Dodson, 1908). Hypothetically, arousal level will vary with perceived question difficulty, and task complexity will vary with text readability. To enable a test of the hypothesis that learning is a complex interactive function

of such text and question characteristics, two sets of inserted post-questions having significantly different difficulty levels (produced by manipulation of distractors for multiple choice items) were varied factorially against a lesson written at three levels of readability (produced by manipulation of lexicon and syntax). A variety of aptitude measures, described below, were gathered and correlated with test results to aid in the interpretation of treatment effects.

A second question of this study concerned the practical utility of conclusions based on previous research on the effects of adjunct questions. Standard control groups have been instructed to read passively. However, observation of the study habits of college students indicated that passive reading is atypical; students may instead be observed to underline, high-light, write notes, and to outline (normative questionnaire data were collected to estimate frequencies). Students in one control group of this experiment were, therefore, directed to study according to their idiosyncratic habits. A second control group was given the traditional direction to read-only.

External validity of research on prose-lesson studying techniques also requires testing of the experimental techniques as they would actually be utilized by students. However, in contemporary research on the use of adjunct questions, the technique which has been employed is one that students would surely disregard if a course grade or graduation were contingent on achievement. Specifically, subjects have been prevented from looking back at the lesson after having encountered the post-question, and even after having attempted to answer it. They have also

been prevented from reviewing the lesson prior to administration of the criterial test. In this study, students were allowed to look back to check their answers to the questions, and they were afforded unrestricted freedom to review the complete lesson prior to testing.

Experiments designed to improve or create new instructional techniques must also employ study materials similar in character to those actually in use, or external validity will be damaged. Furthermore, it may be speculated that improved instructional techniques are most needed for course materials which students find difficult and uninteresting. Given these considerations, an effort was made to select a lesson that students would find boring and hard to study.

METHOD

Lessons

The introduction to a college level text on mathematics was selected. It contained a description of the abstract, deductive character of modern mathematics, a sketch of its history, and the text authors' philosophical perspectives. The reading difficulty level of the text was obviously high. Of a total of 68 sentences, 29 contained 25 or more words. The original text contained 1687 words, and the average number of words for all independent clause units (or simple sentences) was 22.0.

Two versions of the original text were prepared by breaking long sentences into shorter ones, and by replacing uncommon terms with more familiar synonyms. In addition, a concrete example was introduced to clarify an abstract discussion. A moderately revised version contained

1874 words and its independent clause units averaged 18.4 words. A highly revised version contained 1961 words and averaged 14.9 words per clause. Hereafter the original lesson will be referenced as Difficult (D), the moderate revision as Moderately Difficult (M), and the major revision as Average (A). Questionnaire results which support these descriptions will be reported.

The criterial test had been constructed directly from the original lesson. To check the possibility that test information was not adequately preserved or represented in the two revisions, a separate experiment was conducted. Each lesson was studied by independent groups under the expectation of a closed-book test; however, an open-book test was given instead to minimize possible effects of writing style on retention, and to reflect more accurately the available text information.

Adjunct Post-Questions

The lessons were divided into nine, approximately equal, sections (based on results of Frase, 1968) at the same junctures (paragraph boundaries in eight of nine cases), and collated into booklets. A single multiple-choice question stem was written in the knowledge or low-level comprehension domains of the Bloom taxonomy (1956) for each of the nine sections. These questions may also be described by Anderson's categories (1972) as containing one verbatim and four transformed verbatim items, plus four transformed paraphrase items. These questions were specifically, rather than randomly, selected to prevent the

occurrence of any spurious patterns. Two sets of adjunct questions were formed by manipulating the semantic relatedness of the distractors (three). One set of adjunct questions was organized to contain all of the easy items and the other all the hard ones. The adequacy of this manipulation was directly tested by examining item difficulty indices.

The student was instructed to study each lesson page until confident that he could answer a question on its contents. It was explained by the directions that the, "questions are designed to serve you as check points," and that answers were not given for two reasons: "(1) the answers are in the lesson material just read, and you may look back if you have any doubts concerning the correct answer, (2) previous research has shown that most students simply look for the answers provided rather than studying. If you feel reasonably confident about the answer to each question, then you have a sign that your progress is adequate. On the other hand, if you are not sure and have to keep looking back, then you have a sign to slow up and study harder." Although students were permitted to check their answers, they were required to respond on the data sheets before turning back, and they were barred from erasing.

The possibility that the adjunct questions could provide direct, positive transfer to the criterial test, which contained 28 multiple-choice items similar in type to the adjuncts, but not containing any of them, was investigated. A group of 21 students first took the test without any preparation. After this pretest, they were given handouts containing the adjunct question stems and answers and were then re-tested with the handouts as an available aid.

Additional Treatments

Three treatments were employed in addition to relatively hard and easy inserted post-questions: (1) Idiosyncratic study--"Study just as you do typically. If you like to underline, or pen notes in the margins or just read straight, then do so now.": (2) Passive reading--"The method you are to use is the common study technique of reading. You may study in any manner that you please, only do not use your pen or pencil as a study aid."; (3) Underlining--"Please underline potentially significant portions of the text as you would ordinarily do if studying for exams. If you do not ordinarily underline, please try it now."

Underlining was employed only with the D lesson. Line lengths and number of sentences underlined were measured and correlated with test scores and aptitude variables. In addition, biserial correlations were calculated for each test item to determine if underlining of specific text information was predictive of related test performance. To secure reliable judgments identifying correspondence between each test item and the source(s) for the items in the lessons, the investigator and a colleague[2] independently mapped the test to the lesson text. Complete agreement was obtained for 25 of 28 test items; however, four items of the 25 were judged to be represented in more than one text location, and on that basis excluded.

Questionnaire Data

Immediately after studying the experimental lessons, but prior to taking the test, the students were given either a blank sheet, to write any opinions, or a questionnaire which asked about: (a) interest in the

lesson topic; (b) readability level of the text; (c) merit of the experimental study method. Ratings for these questions were secured through use of five point rating scales for which a high (5) rating was most favorable. A second part of the questionnaire asked for information concerning the student's use of active study techniques--underlining, note writing, and outlining--during normal, nonexperimental periods. These data were collected anonymously. Approximately five minutes were expended in distribution, answering, and collecting the questionnaires.

Participants

All students of the undergraduate Educational Psychology course Winter and Spring quarters of 1972 at Southern Illinois University, Carbondale, were required to participate to fulfill a research requirement and the experiment was conducted during regular class hours. To foster academic motivation, students were informed that their instructor would be apprised of any students scoring too low (below chance level) for credit. Data were collected from approximately 700 students, but certain losses occurred. It was discovered after data collection that 15 booklets for the D lesson hard-question treatment were improperly collated. A number of students for whom acquisition data had been collected were absent during the retention testing periods. Attitude data were missing for a small percentage of students, but ACT-E data were missing for about one-third of the sample. To improve the power of significance tests, the largest sample available for each analysis has been utilized. As a consequence, means, standard deviations, and sample sizes are reported for the several analyses. It may be noted later

that the sample size of the Idiosyncratic treatment for the D and M lessons are relatively low; less data had been collected here since the booklets for the other treatments could be re-used. It should also be observed that more data were collected for the D than M lesson, and more data for the M lesson than for the A lesson in accordance with the major interests of this experiment.

Aptitude Measures

Verbal ability scores, American College Testing Program-English subscale (ACT-E), were obtained from administrative records for the majority of students. A variety of attitude measures were obtained by use of a battery administered during the first class meeting in both quarters. The specific scales were: Alpert-Haber Achievement Test Anxiety, Facilitating and Debilitating subscales (1960); Internal-External Locus of Control (Rotter, 1966); Dogmatism (Rokeach, 1960); Social Desirability (Crowne & Marlow, 1964); and Intellectual Self-Confidence, a scale developed by the investigator which is described below.

The Intellectual Self-Confidence Scale (ISCS) is based on a construct defined as follows. Phenomenologically, the belief that one has the capacity to succeed at tasks demanding intellectual effort is the central fact of intellectual self-confidence. Theoretically, the strength and scope of this belief is a function of the individual's reinforcement history. A history of actual, or perceived, success would shape a positive conviction and failure a negative self-regard. In addition, if it is assumed that success constitutes positive rein-

forcement, then cues associated with intellectual activity will themselved acquire reinforcing properties (i.e.. become positive secondary reinforcers). Thus the successful individual will develop a liking for intellectual activity. Furthermore, since the successful individual has been reinforced for his own efforts, self-reliance will also have been shaped. In summary, according to this analysis the belief that one is intellectually capable will necessarily be accompanied by two reinforced attitudes: first, a positive regard for intellectual pursuits; and second, an attitude of self-reliance. Therefore, the construct definition for intellectual-self-confidence incorporates reference to three behavioral tendencies: (1) expectation of success; (2) attraction to intellectual tasks; and (3) self-reliance. The items of the ISCS, 33 in number, have been written to measure one or more of these three component tendencies. The ISCS's reliability is approximately .75, Cronbach alpha. Validation studies are reported in Kirby & Hiller, 1973; and Hiller, 1972.

Procedures

The goal of this experiment, to develop information on the effectiveness of different study techniques, was described on the cover of all lesson booklets. The treatments were not, however, described to the classes. An attempt to strengthen external validity was made by asking the students to, "work about as hard as you normally do. You should not make any special effort to do well, nor should you give up because a grade is not at stake. The worth of the recommendations we can make to you and other students are obviously based completely on

your cooperation." The final paragraph of the cover stated that there would be 25 minutes study time and that the lesson contained about 1800 words--"This means that you will have time to read it twice over--in other words you have enough time to study." (Only a negligible number of students were observed to be studying actively when the 25 minute period ended, so that unavailability of time was not a factor. Questionnaire responses and the serial learning curves, to be reported, further demonstrated that lack of study time did not differentially effect study behavior.) The students were also informed that the test was designed to, "test your retention of facts and your understanding of the ideas presented in the lesson."

The second page of all booklets described the student's study technique. In summary, the order of events was as follows:

(1) The students read the orienting directions printed on the lesson cover sheets;

(2) Next they read the treatment direction on the second page, and then studied for 25 minutes;

(3) The lessons were collected and questionnaires were distributed;

(4) After five minutes, the questionnaires were collected, the test distributed, and the students were given 20 minutes testing time.

Retention

In addition to the test for acquisition, two weeks later the identical criterial test was administered without warning to all treatment groups having studied the D lesson to obtain an estimate of

retention effects. Students having studied the M and A lessons were given ten minutes to re-study their lesson booklets prior to re-taking the test; this procedure was designed to explore the possibility that note-writing and underlining of lesson material facilitates review activities, and to improve external validity, on the assumption that most students will study or review before taking a test.

RESULTS

Acquisition and Retention

Analysis of acquisition test results (see Table 1, part A) demonstrated that Idiosyncratic study and Passive Reading were similar to each other, but superior to both inserted question treatments. Overall ANOVA for the experiment (see Table 2, part A) determined that treatment effects were highly significant.[3] Lesson effects also were found to be significant, at $p < .002$. Inspection of results across the lessons (Table 1) shows that acquisition increased from the D lesson ($M = 12.5$, $SD = 3.1$) to the M lesson ($M = 13.5$, $SD = 3.8$) but that acquisition for the A lesson ($M = 13.5$, $SD = 3.3$) did not increase further (see Figure 1).

- - - - - - - - - - - - - - - - - -

Insert Tables 1, 2
and Figure 1 about here

- - - - - - - - - - - - - - - - - -

The absence of a difference between performances for the two revised lessons might have reflected a diminished correspondence between the test and lesson contents. However, the following results for the groups given the open book test imply that test relevant information had been adequately preserved and that the text revisions were effective:

(a) Difficult Lesson M = 15.9, SD = 3.5, N = 11;

(b) Moderately Difficult M = 17.6, SD = 2.5, N = 18;

(c) Average M = 19.5, SD = 2.65, N = 11.

These results are significantly different, $F(2,37) = 4.63$, $p < .03$. In addition, the difference between results for the M and A lessons is significant, $t(27) = 2.07$, $p < .05$ under a nondirectional test.

Retention data (Tables 1B and 2B) demonstrated no significant effects. Since results for the immediate test indicate that the inserted questions lowered the acquisition of information incidental to the questions, and that the two control groups learned about equally well, retention data have also been analyzed separately for the control and experimental groups, using the two revised lessons. Results for the inserted question groups did not demonstrate significant differences in retention, but the controls did differ significantly across the M and A lessons; retention was higher for the M lesson, $t(150) = 2.63$, $p < .003$. Parallel to this result, the controls did not show a significant loss of retention in the M lesson, while controls using the A lesson did drop.

The Easy inserted questions achieved an average difficulty level across the three lessons of 69%, while the Hard questions averaged 53% (Table 3). The difference between the two sets of questions was highly significant.

Insert Table 3 about here

The group tested to determine if the inserted questions provided direct positive transfer to the test averaged 7.5 on the pre-test and 7.9 on the post-test, $\underline{t}$ (20) = .24.

Examination of Figure 1 suggests that treatments interacted with lessons for the inserted question conditions. The ANOVA does indicate a tendency toward interaction at $p < .06$:

LESSONS F (2,323) = 2.97 $p < .05$

TREATMENTS F (1,323) = .60 ns

INTERACTION F (2,321) = 2.78 $p < .06$

Theoretically, inserted questions enhance incidental learning by helping to stimulate and maintain study activity. Table 4, and Figures 2, 3, and 4 display results showing how students performed on the test as a function of the ordinal sequence of test relevant information presented by the lessons. Each point in the Figures is the average of four test items. In the D lesson, it may be seen that acquisition declined rapidly, although, it rose somewhat toward the end of the lesson for

- - - - - - - - - - - - - - - - -

Insert Table 4 and
Figure 2, 3 and 4 about here

- - - - - - - - - - - - - - - - -

Passive reading and inserted question groups. Performance by the inserted question groups does not appear to have declined as rapidly in the two control groups, but was lower at the outset. Performance for the M lesson was similar to that for the D lesson. The major difference between results for the A lesson and the other two was that relatively

high performance was maintained further into the text for both control groups and Easy questions.

There had been 21 test items available for the serial position analysis presented above. One item which had not been included in the analyses (presented in Table 4, and Figures 2, 3, and 4) is interesting for the fact that it was recognizably different from the others. This item pertained to the definition of "metaphysics" which was said to be "(a basic branch of philosophy which may be called the philosophy of philosophy)," and the multiple-choice test item had as the correct response, "philosophy of philosophy." This item of information occurred toward the end of the seventh page of the nine page text, and falls within the fourth set of test items shown in Figures 2, 3, and 4. Average results for the fourth sets without the metaphysics item were 38%, 40%, and 36% for the D, M, and A lessons respectively. In contrast to these low values, performances for the metaphysics item were 80%, 77%, and 57%, thereby suggesting a Von Restorff effect. This result strongly implies that most students had read the lesson at the item's text location despite the fact that performance for surrounding learning was only slightly above chance (25%). Furthermore, this result cannot be attributed to guessing since in the control group, which had not read the lesson, only one student of 21 chose the correct response.

Aptitude-Treatment Correlations and Interactions

Aptitude-acquisition test correlations are shown in Table 5. The measure of intellectual confidence (ISCS) correlated significantly with acquisition performance for Passive Reading, and with that of inserted

question treatments in the D lesson, and was most highly correlated with performance for the Hard question treatment. In the M lesson, ISCS correlated only with scores for the Hard question group, while it was uncorrelated with any treatment results in the A lesson. The facilitating text anxiety measure correlated with both control group results and with those for Easy questions in the D lesson, but was not correlated with any other treatment conditions. Debilitating anxiety produced correlations that mirrored those for the ISCS. Verbal ability was generally correlated with performance for the various treatments, and showed a tendency in all three lessons to be more predictive of performance for Passive reading than Idiosyncratic study.

Insert Table 5 about here

Results for Dogmatism, S.D., and I-E scales were not generally significant and therefore are not reported here.

Retention-aptitude correlations for the identical students employed in the acquisition analyses (Table 5) are displayed in Table 6. Here ISCS emerged as a significant correlate for the Idiosyncratic treatment of the D lessons, while the correlation for Hard questions dropped from .43 to .26. In the M lesson, Passive reading and Easy questions were raised to significant values while the correlation for Hard questions dropped from .41 to .18. Facilitating anxiety maintained significant correlation in the D lesson only for Passive reading, but rose to significance for the Idiosyncratic study treatment of the M lesson. Verbal ability demonstrated improved prediction for the revised lessons, but overall slight losses for the D lesson.

- - - - - - - - - - - - - - - - - - -

Insert Table 6 about here

- - - - - - - - - - - - - - - - - - -

Data were graphed and examined for disordinal aptitude-treatment interactions. Non-significant, but consistent patterns indicated that low aptitude students performed best under Idiosyncratic directions, while high aptitude students performed best under Passive reading.

Questionnaire Findings

Student ratings for lesson readability (Table 7) demonstrate that the text manipulations were effective. (The rating scale description corresponding to the value "2" was "rather difficult;" while "3" corresponded to, "about average for textbooks.")

- - - - - - - - - - - - - - - - - - -

Insert Table 7 about here

- - - - - - - - - - - - - - - - - - -

The students indicated clear differences in their approval for the study methods (Table 8). Passive reading was consistently the least preferred study technique, with Hard and Easy questions less preferred than study according to habit, for the D and M lessons. For the A lesson, only Passive reading received a negative rating (value "2" asserted, I rather dislike the method and probably would not use it; "3" indicated neutrality; and "4" reversed "2").

Insert Table 8 about here

The ratings for topic liking (Table 9) show that interest on the average varied from "dislike somewhat," at scale value "2" to "neutral" at "3". Although treatments did not affect these ratings,

lessons did. It may be seen that interest was positively related to lesson readability.

- - - - - - - - - - - - - - - - -

Insert Table 9 about here

- - - - - - - - - - - - - - - - -

Normative data for active study habits (Table 10) showed that the students sampled are typically quite active when studying. Eighty-eight percent indicated that they sometimes or usually write notes while studying. It should be noted, however, that only 17% of the students in the Idiosyncratic (or habit) groups wrote notes on their lesson booklets. Eighty-four percent indicated that they sometimes or usually underline, and 73% of the students in the habit groups did underline in this experiment.

- - - - - - - - - - - - - - - - -

Insert Table 10 about here

- - - - - - - - - - - - - - - - -

Underlining Performance Analyses

Analysis of results for the underlining treatment did not determine any significant results nor tendencies. Underlined portions of text did not yield prediction of scores for corresponding test items. Furthermore, the quantitative measures of underlining were not correlated with any of the aptitude measures.

DISCUSSION

In contrast to generalizations drawn from prior research, the use of inserted post-questions, with review permitted, tended to lower incidental

learning. The serial acquisition curves shown in Figs. 2, 3, and 4 clearly demonstrate that the inserted questions did not serve to maintain learning throughout the lessons. Instead, the questions may have diverted attention from information incidentally related to the question topics. Kulhavy (1972) obtained comparable negative results regarding incidental learning and drew a similar conclusion from an experiment in which the adjunct study aid took the form of an embedded sentence having information marked out with black ink that was presented elsewhere in the same paragraph.

The fact that the self-confidence and debilitating anxiety scales correlated with learning for the inserted question groups studying the D lesson, and also for the Hard question group having the M lesson, implies that distracting emotional behaviors were provoked (see Wine, 1971). Since answers were not provided for the inserted questions, performance for anxious students may have been depressed, based on the finding of Campeau (1968) that acquisition and retention were significantly lower for highly anxious females who were deprived of feedback while learning from programmed text. Students with initially low confidence (expectation of success) may also have tended to reduce effort when finding their performance low, whereas students with high confidence may have worked harder (Means & Means, 1971).

Since confidence and anxiety measures did not correlate with learning in the A lesson, the inferior performance of the Hard question group here may be solely attributed to diverted attention.

Consistent with the extensive survey by Bracht (1970), disordinal aptitude-treatment interactions were not significant.

The result that student-initiated underlining failed to raise learning is consistent with previous research by Arnold (1942), Stordahl and Christensen (1956), and Idstein and Jenkins (1972). One aspect of the underlining group's performance was the sheer quantity of the underlining. On the average, students underlined 23.5 sentences, or 34.5% of all sentences, and this level was significantly above that of students in the Idiosyncratic group who averaged 15.1, $\underline{t}$ (116) = 4.0, $p < .005$. Additionally, the underlining group averaged 210 inches of underlining as compared to 102 inches for the habit group, $\underline{t}$ (116) = 4.8, $p < .005$. This rather excessive underlining activity may have served to interfere with concentration on a manageable amount of information, and some students wrote that they found themselves so busy underlining that comprehension of the lesson was interfered with.

The lack of correlation between underlining of text and performance on associated test items may also be accounted for, in part, by assuming that there are at lest two reasons for underlining: (1) establishing cues for later review of significant material, and (2) identification of difficult material which requires more careful study for comprehension. Despite the fact that instructions to underline did not raise learning, the students indicated on the questionnaires their belief that attempting to underline while studying usually helps to maintain attention ("stay awake").

It had been expected that Idiosyncratic study would prove superior to Passive reading. Although Passive reading did yield stronger correlations than Idiosyncratic study for confidence, debilitating anxiety, and verbal ability with acquisition scores, mean group performances for the two methods failed to differ significantly. A possible

explanation, in addition to counterbalanced motivational effects, concerns the nature of the experimental lessons' contents and organization, quite apart from factors of sentence length and vocabulary. Frase (1969) found that students spontaneously wrote notes when studying text that placed a large strain on immediate memory; however, it was the case here that only about 17% of the students took notes (as compared to the questionnaire report according to which 88% said that they usually or sometimes write notes), and it was observed that note writing typically extinguished by the third of the nine lesson pages. Another point in addition to the fact that note writing was relatively infrequent, is the finding by Schultz and Di Vesta (1972) that taking notes was initially advantageous where the text presented an unusual structure (information organized according to attributes rather than names). Both on empirical and logical grounds, it would seem that control subjects should be permitted to study according to individual preference or habit.

One potentially interesting result was that retention for the M lesson was superior to that for the A lesson, despite the fact that comprehension scores and readability ratings for the A lesson were higher. It would be tempting to infer that the added difficulty of the M lesson successfully stimulated careful reading, and that texts should therefore be written to require effort beyond some minimally comfortable level. However, as Rothkopf (1970) has pointed out, an important pedagogical consideration is the student's placing of himself in a learning situation, and many students may habitually delay and avoid study of materials perceived as difficult. Strong support for such an idea was

obtained by Klare and Smart (in press) who found a correlation of .75 between readability and the probability of course completion for the U. S. Armed Forces correspondence courses (with course length controlled).

It is important to recognize the constraints on the use of inserted questions in this experiment. The adjunct questions may have tended to lower incidental learning because of the technique's novelty. It is a possibility that continued practice with questions in applied settings would reverse the present results. In addition, the questions dealt with facts or narrow topics rather than with overarching ideas. Some students complained about this and suggested that the questions would have been more useful had they been aimed at the "main ideas." The finding of Watts and Anderson (1971), that questions which required the application of principles were more effective than questions which required recognition of previously described examples, supports this proposal. It is interesting to note that Watts and Anderson raised the possibility that their application questions were more effective simply because they were measurably harder; however, results for the Hard question groups in this experiment imply that additional processing demands rather than difficulty per se explain their results.

It should also be clearly recognized that the lesson material used here was more nearly a collection of pieces of discrete information rather than being tightly woven or sequentially dependent. It seems reasonable to suppose that questions which enable a student to check his progression through highly integrated materials, such as mathematical

text, would be very valuable, particularly if the questions were "strategically" placed rather than inserted at "frequent" intervals.

The results for both acquisition and two-week retention caution against reliance on low level cognitive questions inserted in text as a general technique for promoting the learning of information not specifically cued. Furthermore, there yet remains unanswered a most important practical question--how will students actually use inserted questions when studying for school examinations? If the inserted questions were not well represented in the examinations, we may predict that attention to them would extinguish. On the other hand, extensive representation would encourage careful attention to the questions (or surrogate instructional objectives) during study. Unfortunately, we might then also predict that students would look ahead to find such valuable test clues when studying, and thereby risk the loss of incidental learning found by Rothkopf (1966) and others for pre-question treatments.

References

Alpert, R., & Haber, N. Anxiety in academic achievement situations. Journal of Abnormal and Social Psychology, 1960, 61, 207-215.

Anderson, R. How to construct achievement tests to assess comprehension. Review of Educational Research, 1972, 42, 145-170.

Arnold, H. F. The comparative effectiveness of certain study techniques in the field of history. Journal of Educational Psychology, 1942, 33, 449-457.

Bloom, B. S. (Ed.) Taxonomy of educational objectives. I: Cognitive Domain. New York: David McKay Company, 1956.

Bracht, G. Experimental factors related to aptitude-treatment interactions. Review of Educational Research, 1970, 40, 627-645.

Campeau, P. L. Test anxiety and feedback in programmed instruction. Journal of Educational Psychology, 1968, 59, 159-163.

Crowne, D., & Marlow, D. The approval motive. New York: John Wiley & Sons, 1964.

Frase, L. T. Effect of question location, pacing, and mode upon retention of prose material. Journal of Educational Psychology, 1968, 59, 244-249.

Frase, L. T. Cybernetic control of memory while reading connected discourse. Journal of Educational Psychology, 1969, 60, 49-55.

Frase, L. T. Boundary conditions for mathemagenic behavior. Review of Educational Research, 1970, 40, 337-348.

Hiller, J. H. "Effects of individual differences on the learning and retention of ordinary prose text." Mimeographed paper, 1972, results to be reported in Aptitudes and Instructional Methods, L. J. Cronbach and R. Snow (in press).

Idstein, P., & Jenkins, J. R. Underlining versus repetitive reading. Journal of Educational Research, 1972, 65, 321-323.

Kirby, E. A., & Hiller, J. H. "Comparative validation of a direct and an indirect measure of academic self-confidence." Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, February, 1973.

Klare, G. R. & Smart, K. L. Analysis of the readability level of selected United States Armed Forces Institute Printed Instructional Materials. Journal of Educational Research, (in press).

Kulhavy, Raymond W. Effects of embedding orienting stimuli in a prose passage. Psychonomic Science, 1972, 28, 213-4.

McGaw, Barry & Grotelueschen, Arden. Direction of the effect of questions in prose material. Journal of Educational Psychology, 1972, 63, 580-588.

Means, R. S., & Means, G. S. Achievement as a function of the presence of prior information concerning aptitude. Journal of Educational Psychology, 1971, 62, 185-187.

Quellmalz, E. "Effects of response requirement type, placement, and category upon attention patterns influencing prose learning." Paper presented at the Meeting of the American Educational Research Association, Chicago, April, 1972.

Rokeach, M. The open and closed mind. New York: Basic Books, 1960.

Rothkopf, E. Z. Learning from written materials: an exploration of the control of inspection behavior by test-like events. American Educational Research Journal, 1966, 3, 241-249.

Rothkopf, E. Z. The concept of mathemagenic activities. Review of Educational Research, 1970, 40, 325-336.

Rothkopf, E. Z. & Bisbicos, E. E. Selective facilitative effects of interspersed questions of learning from written materials. Journal of Educational Psychology, 1967, 58, 56-61.

Rotter, J. Generalized expectancies for internal versus external control of reinforcement. Psychological Monographs, 1966, 80, (1, Whole No. 609).

Schultz, C., & Di Vesta, F. Effects of passage organization and note taking on the selection of clustering strategies and on recall of textual materials. Journal of Educational Psychology, 1972, 63, 244-252.

Stordahl, K. E., & Christensen, C. H. The effect of studying techniques on comprehension and retention. Journal of Educational Research, 1956, 49, 561-570.

Watts, G., & Anderson, R. Effects of three types of inserted questions on learning from prose. Journal of Educational Psychology, 1971, 62, 387-396.

Wine, J. Test anxiety and direction of attention. Psychological Bulletin, 1971, 76, 92-104.

Yerkes, R. M., & Dodson, J. D. The relations of strength of stimulus to rapidity of habit-formation. Journal of Comparative Neurology, 1908, 18, 459-482.

Footnotes

[1]This paper was originally presented at the Meeting of the American Psychological Association, Division 15, Honolulu, Hawaii, September, 1972. The data were collected while the author was at Southern Illinois University, Carbondale. The author wishes to acknowledge his indebtedness to E. Z. Rothkopf and L. T. Frase for their 1970, AERA Presession which stimulated this research, and to thank George Klare whose own research on readability provided the model.

[2]Reed Williams performed this labor and also double checked all experimental materials, as well as providing the opportunity for a pilot test, for which help the author is very grateful.

[3]The ANOVA was performed by the multiple regression technique and program described in Research Design in the Behavioral Sciences, Multiple Regression Approach. F. J. Kelly, D. L. Beggs, and K. A. McNeil.

TABLE I

Test Results

A. ACQUISITION

TREATMENTS	LESSONS: LOW READABILITY M	SD	N	MODERATE READABILITY M	SD	N	AVERAGE READABILITY M	SD	N
1. Passive Reading	13.1	3.3	86	14.4	3.9	68	13.7	3.2	35
2. Idiosyncratic	13.3	2.1	59	14.4	3.7	41	14.2	3.3	53
3. Easy questions	12.1	3.1	81	12.2	3.3	60	13.4	3.3	34
4. Hard questions	11.4	3.1	64	13.1	3.9	55	12.2	3.0	33
5. Underlining	12.8	2.7	59						

B. RETENTION

TREATMENTS	LESSONS: LOW READABILITY M	SD	N	MODERATE READABILITY M	SD	N	AVERAGE READABILITY M	SD	N
1. Passive Reading	11.1	3.1	71	13.7	4.9	54	11.6	4.1	27
2. Idiosyncratic	11.1	2.4	47	13.7	4.0	33	11.9	3.2	38
3. Easy questions	10.6	3.3	69	11.9	3.8	47	11.4	3.5	24
4. Hard questions	10.2	2.9	53	12.9	4.3	46	12.2	3.3	25
5. Underlining	11.1	2.4	47						

TABLE 2

ANOVA For Acquisition and Retention

A. Acqusition, Excluding Underlining Treatment		
1. All lessons		
LESSONS	$F(2,663) = 6.86$	$p < .002$
TREATMENTS	$F(3,663) = 11.57$	$p < 1.0 \times 10^{-7}$
INTERACTION	$F(6,657) = 1.55$	ns
2. Difficult lesson		
TREATMENTS	$F(3,286) = 6.09$	$p < .0005$
3. Moderate difficulty lesson		
TREATMENTS	$F(3,220) = 4.74$	$p < .003$
4. Average difficulty lesson		
TREATMENTS	$F(3,151) = 2.92$	$p < .036$
B. Retention, Excluding Underlining Treatment		
1. Difficult lesson		
TREATMENTS	$F(3,236) < 1$	
2. Moderate and Average difficult lessons		
No effects were significant at		$p < .05$

TABLE 3

Inserted Question Results*

TREATMENTS	LESSONS								
	LOW READABILITY			MODERATE READABILITY			AVERAGE READABILITY		
	M	SD	N	M	SD	N	M	SD	N
1. Easy questions	6.48	1.4	81	5.88	1.8	60	6.09	1.2	34
2. Hard questions	4.66	1.7	64	4.84	1.6	55	4.67	1.8	33

*Note: Lessons $F(2,323) = .78$ ns
Treatments $F(1,323) = 65.31$ p 1.0×10^{-7}
Interaction $F(2,321) = 1.82$ ns

TABLE 4

Mean Percent Correct for Five Sets of Test Items (four items per set) Grouped by Information Sequence

LESSON			MEAN PERCENT CORRECT ARRANGED BY SEQUENCE				
DIFF.	TREATMENTS	N	1 M%	2 M%	3 M%	4 M%	5 M%
	1. Passive Reading	86	70	56	41	39	47
	2. Idiosyncratic	57	77	62	51	39	40
	3. Easy Questions	83	61	49	44	40	48
	4. Hard Questions	63	59	54	35	37	48
	5. Underlining	63	68	56	45	36	41
MOD.	1. Passive Reading	68	71	63	50	42	55
	2. Idiosyncratic	42	74	63	51	36	51
	3. Easy Questions	60	65	46	34	40	50
	4. Hard Questions	55	66	55	48	42	50
AVER.	1. Passive Reading	36	71	67	45	38	57
	2. Idiosyncratic	52	70	69	45	31	51
	3. Easy Questions	36	63	62	49	40	44
	4. Hard Questions	33	59	55	44	34	47

TABLE 5

Acquisition Test Results with Test - Aptitude Correlations

LESSON	TREATMENTS	TEST RESULTS			TEST-APTITUDE CORRELATIONS*				
		M	SD	N	ISC	FA	DA	ACT-E	(N)
DIFF.	1. Passive Reading	13.0	3.4	63	30*	32*	-25*	46*	(50)
	2. Idiosyncratic	13.6	1.9	44	-02	32*	-10	33*	(43)
	3. Easy Questions	12.0	3.3	61	31*	-04	-21	60*	(52)
	4. Hard Questions	11.0	3.0	51	43*	33*	-43*	35*	(41)
	5. Underlining	12.8	2.3	43	26	-21	-19	21	(33)
MOD.	1. Passive Reading	14.5	4.3	51	21	10	-20	43*	(44)
	2. Idiosyncratic	14.3	4.0	30	13	22	-23	33	(30)
	3. Easy Questions	12.3	3.5	38	12	-11	07	18	(35)
	4. Hard Questions	13.1	4.1	44	41*	11	-45*	29	(42)
AVER.	1. Passive Reading	13.8	3.2	25	09	13	07	41	(21)
	2. Idiosyncratic	14.4	3.0	34	12	12	-25	22	(31)
	3. Easy Questions	13.5	3.2	22	26	06	-14	50	(15)
	4. Hard Questions	12.9	2.7	21	28	05	-29	-12	(16)

*NOTE $p. < .05$ non- directional test

TABLE. 6

Retention Test Results with Test - Aptitude Correlations

LESSON	TREATMENTS	TEST RESULTS			TEST-APTITUDE CORRELATIONS*				
		M	SD	N	ISC	FA	DA	ACT-E	(N)
		11.0	3.1	63	32*	26*	-18	28	(43)
DIFF.	1. Passive Reading	11.1	2.5	44	38*	06	-11	25	(38)
	2. Idiosyncratic	10.6	3.4	61	17	02	-13	42*	(45)
	3. Easy Questions	10.3	2.8	51	26*	23	-41*	55*	(37)
	4. Hard Questions	10.9	2.3	43	-03	-03	05	09	(31)
	5. Underlining								
MOD.	1. Passive Reading	13.8	5.0	51	30*	17	-05	53*	(37)
	2. Idiosyncratic	13.3	3.9	30	33	39*	-26	56*	(23)
	3. Easy Questions	11.4	3.7	38	37*	11	05	36	(27)
	4. Hard Questions	12.7	4.3	44	18	01	-28	33*	(36)
AVER.	1. Passive Reading	11.3	4.0	25	-11	31	-14	53*	(21)
	2. Idiosyncratic	12.2	3.2	34	-06	14	-15	28	(31)
	3. Easy Questions	11.8	3.4	22	00	38	-24	63*	(15)
	4. Hard Questions	12.6	3.4	21	27	-02	-05	24	(16)

*NOTE $p. < .05$ non-directional test

TABLE 7

Questionnaire Ratings For Lesson Readability*

TREATMENTS	LESSONS								
	DIFFICULT			MODERATE			AVERAGE		
	M	SD	N	M	SD	N	M	SD	N
1. Passive Reading	2.3	.66	29	2.5	.64	37	2.8	.86	28
2. Idiosyncratic	2.0	.84	30	2.8	.59	37	3.1	.89	45
3. Easy Questions	2.2	.80	30	2.5	.80	36	3.0	.96	35
4. Hard Questions	2.2	.60	22	2.5	.59	42	2.7	.58	34

* Note: Lessons $F(2,400) = 20.56$ $p < 1.0 \times 10^{-7}$
Treatments $F(3,400) = 2.00$ ns
Interaction $F(6,394) < 1.$

TABLE 8

Questionnaire Ratings For Study Technique*

TREATMENTS	LESSONS								
	DIFFICULT			MODERATE			AVERAGE		
	M	SD	N	M	SD	N	M	SD	N
1. Passive Reading	2.8	1.4	29	2.8	1.1	37	2.8	1.2	28
2. Idiosyncratic	3.9	1.3	30	3.9	1.0	37	3.6	1.1	45
3. Easy Questions	3.3	1.0	30	3.2	1.1	36	3.5	.9	35
4. Hard Questions	2.9	1.1	22	3.1	1.0	42	3.5	1.2	34

* Note: Lessons $F(2,400) < 1.$
Treatments $F(3,400) = 12.52$ $p < 1.0 \times 10^{-7}$
Interaction $F(6,394) < 1.$

TABLE 9

Questionnaire Ratings for Topic Liking*

TREATMENTS	LESSONS								
	DIFFICULT			MODERATE			AVERAGE		
	M	SD	N	M	SD	N	M	SD	N
1. Passive Reading	2.2	1.2	29	2.1	1.0	37	2.5	1.3	28
2. Idiosyncratic	1.8	1.2	30	2.6	1.1	37	2.4	1.0	45
3. Easy Questions	2.0	1.0	30	2.0	.9	36	2.6	1.3	35
4. Hard Questions	2.0	1.1	22	2.4	1.2	42	2.6	1.0	34

*Note: Lessons $F(2,400) = 4.68$ $p < .01$
Treatments $F(3,400) < 1$
Interaction $F(6,394) = 1.28$ ns

TABLE 10

Self-Reported Frequencies For Active Study Techniques (N=428)

STUDY TECHNIQUE	USAGE FREQUENCY		
	% USUALLY	% SOMETIMES	% NEVER
UNDERLINING	47	37	16
TAKING NOTES	47	41	12
OUTLINING	15	35	50

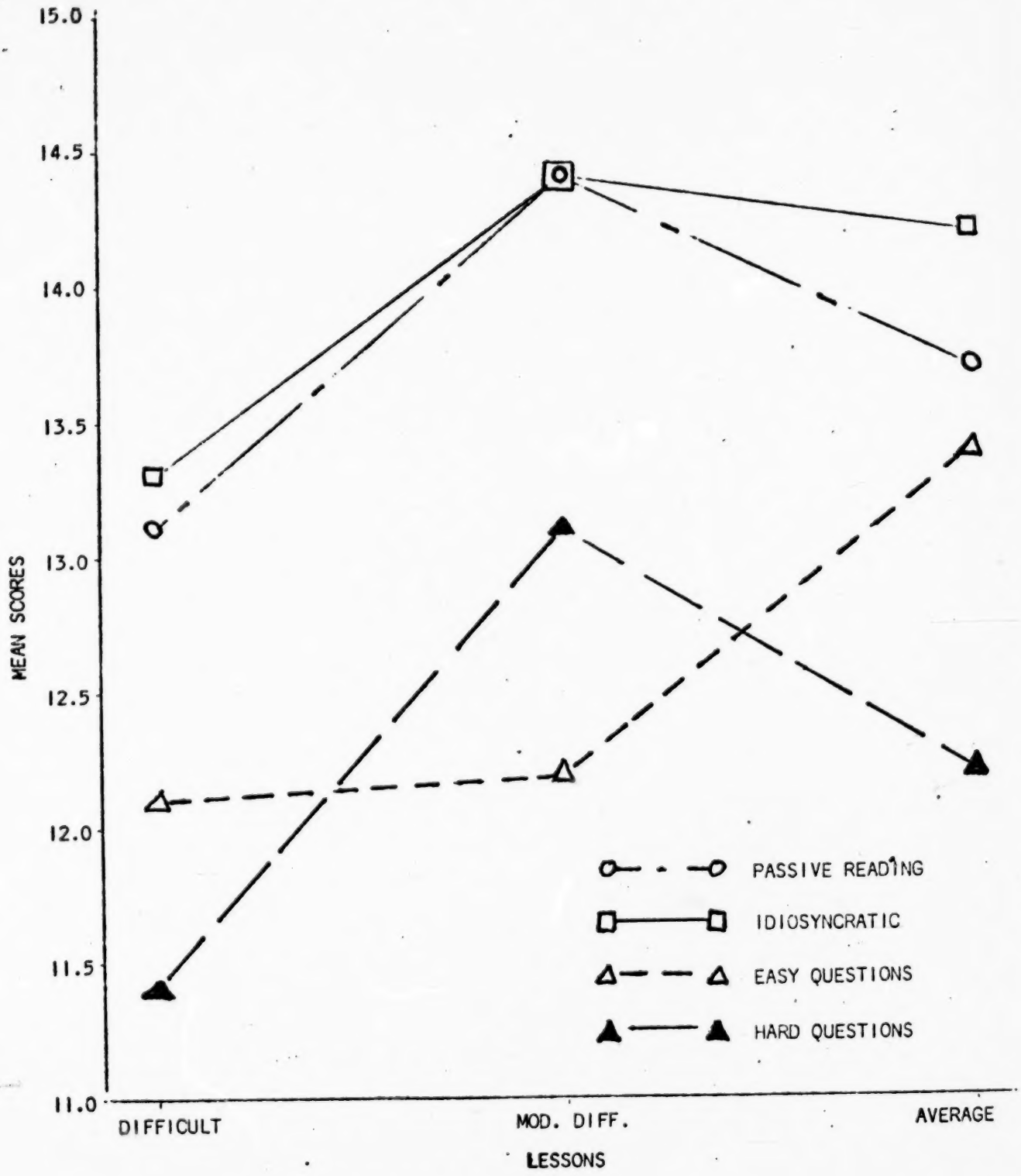

Fig. 1. Mean acquisition test scores.

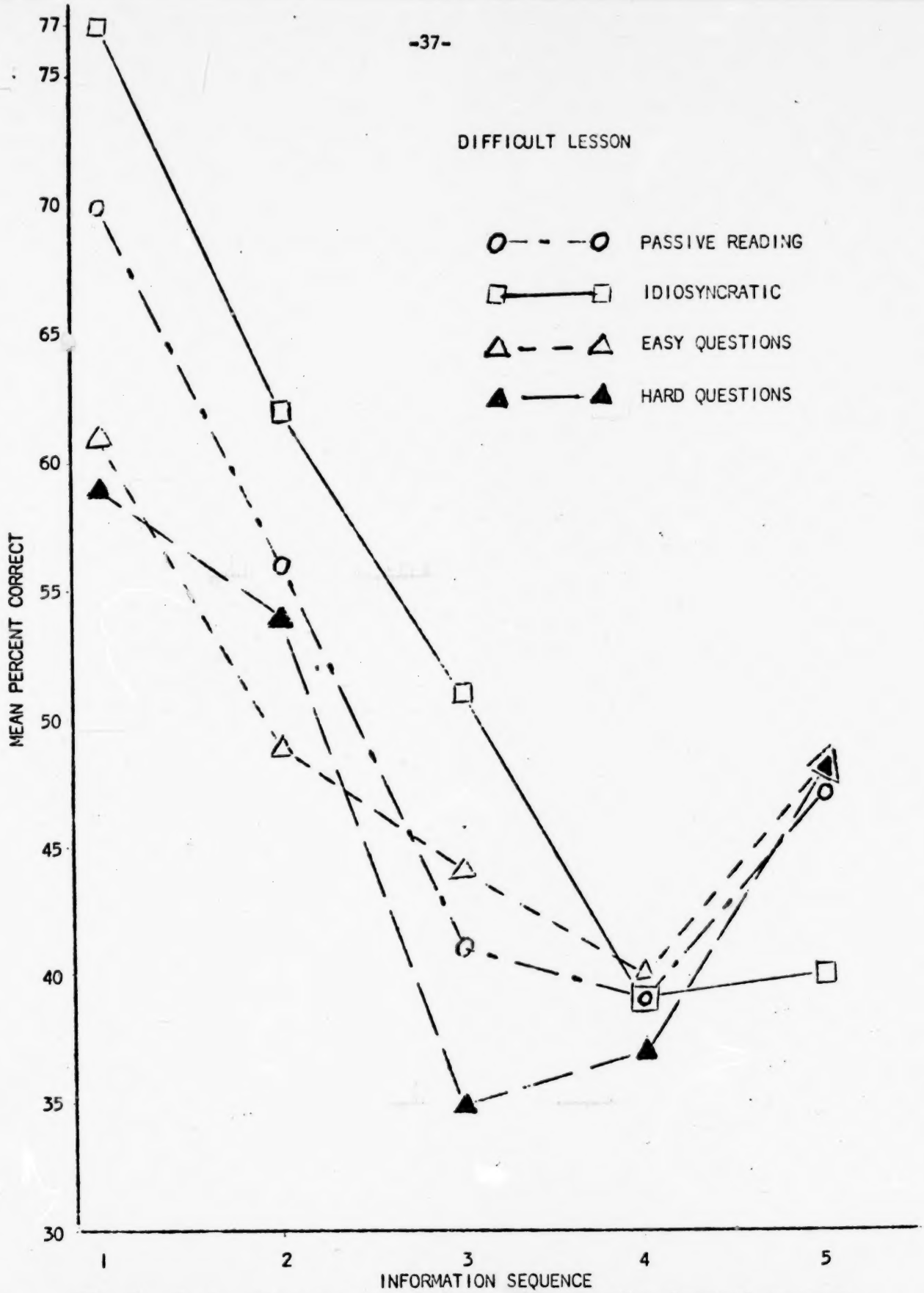

Fig. 2. Test performance as a function of information (answer) sequence for the Difficult lesson.

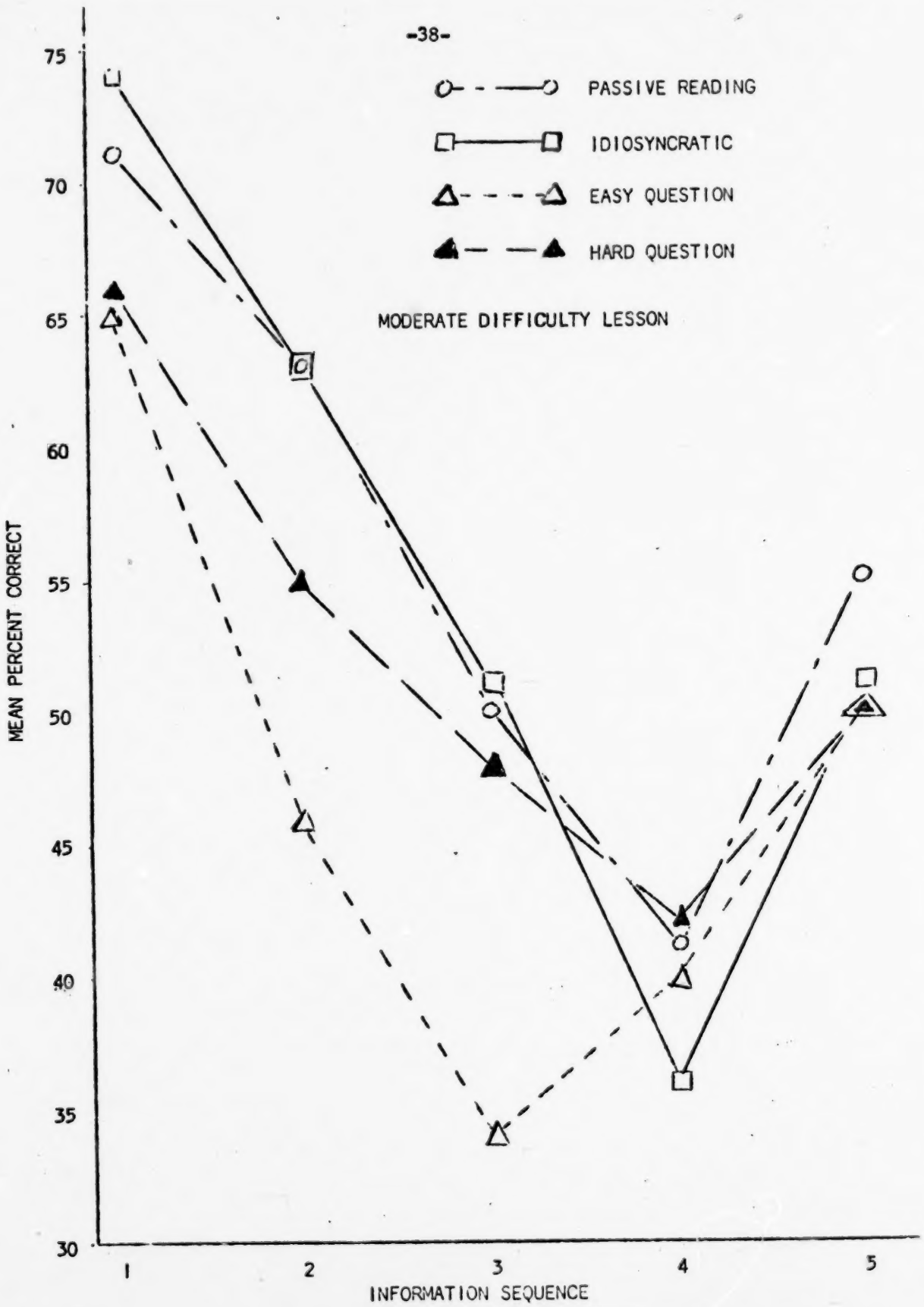

Fig. 3. Test performance as a function of information (answer) sequence, for the Moderate Difficulty lesson.

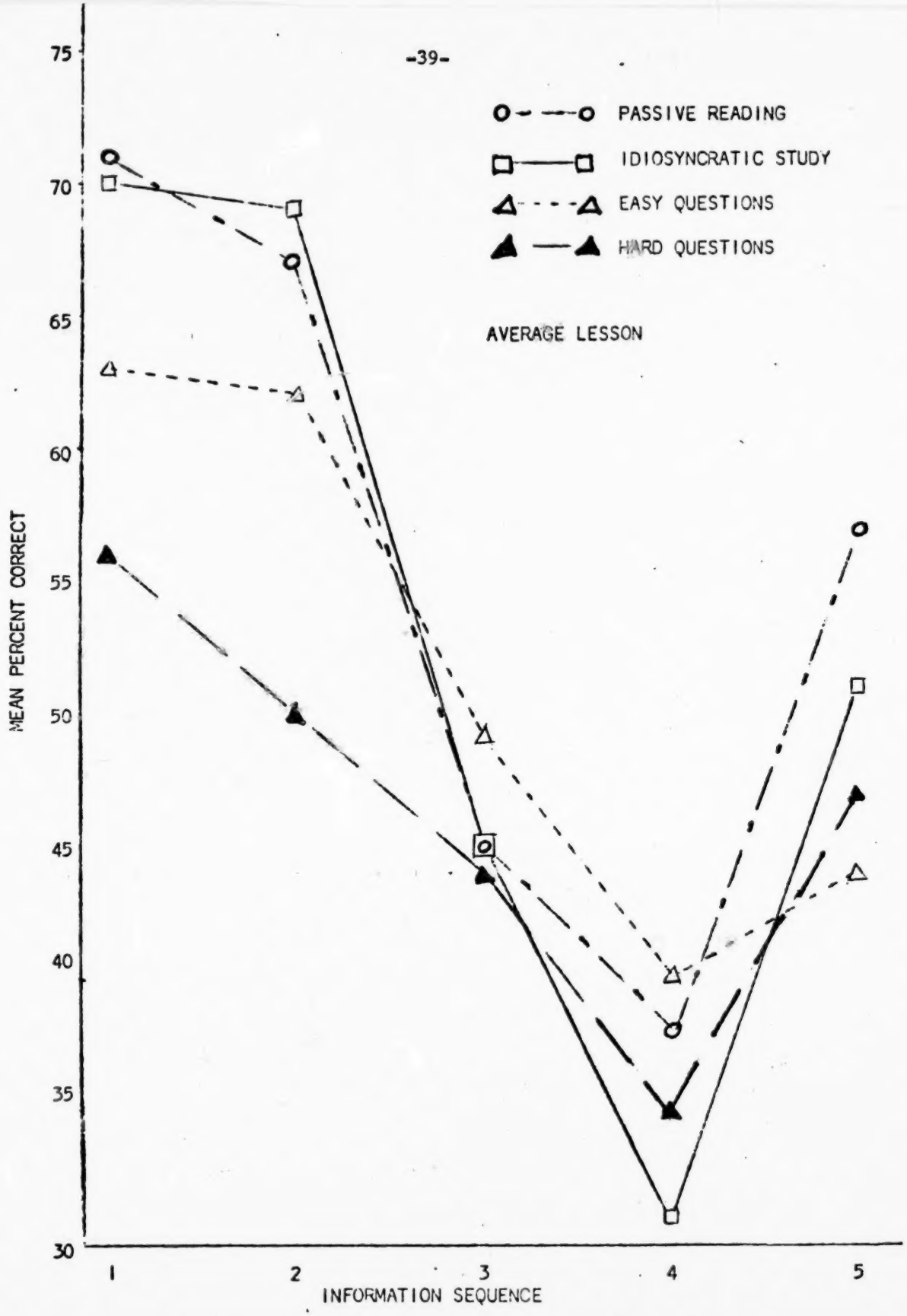

Fig. 4. Test performance as a function of information (answer) sequence, for the Average lesson.